



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2013

Four novel papillomavirus sequences support a broad diversity among equine papillomaviruses

Lange, C E ; Vetsch, E ; Ackermann, M ; Favrot, C ; Tobler, K

Abstract: Papillomaviruses appear to be species-specific pathogens, and it was suggested that each animal species might harbour its own set of papillomaviruses. However, all approaches addressing the underlying evolutionary phenomena still suffer from very limited data about animal papillomaviruses. In case of the horse for example, only three equine papillomaviruses (EcPVs) have been identified. To further address the situation in this host, suspected papillomavirus-associated lesions were tested for EcPV DNA. Four novel EcPV types were detected and their genomes entirely cloned and sequenced. They display the characteristic organization, with early (E) and late (L) regions harbouring the seven classical open reading frames divided by non-coding regions. They were named EcPVs 4, 5, 6 and 7, according to their dissimilarity to other papillomaviruses. Most L1 nucleotide identities were shared with EcPV2 in case of EcPV4 (62 %) and EcPV5 (60 %) or with EcPV3 in case of EcPV6 (70 %) and EcPV7 (71 %). Thus, EcPVs 4 and 5 may establish novel species within the genus Dyoiota, while EcPVs 6 and 7 might fit into the genus Dyorho and belong to the same species as EcPV3. They were found in genital plaques (EcPV4), aural plaques (EcPV5, EcPV6) or penile masses (EcPV7). Interestingly, PCR analysis revealed the DNA of EcPV2 and EcPV4 as well as of EcPV3 and EcPV6 together in the same tissue samples, respectively. In conclusion, the DNA of four novel EcPV types was identified and cloned. They cluster with the known types and support broad genetic EcPV diversity in at least two of the known clades. Furthermore, PCR assays also provide evidence for EcPV co-infections in horses.

DOI: <https://doi.org/10.1099/vir.0.052092-0>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-79401>

Journal Article

Accepted Version

Originally published at:

Lange, C E; Vetsch, E; Ackermann, M; Favrot, C; Tobler, K (2013). Four novel papillomavirus sequences support a broad diversity among equine papillomaviruses. *Journal of General Virology*, 94(Pt 6):1365-1372.

DOI: <https://doi.org/10.1099/vir.0.052092-0>

Four novel papillomavirus sequences support a broad diversity among equine
papillomaviruses

Running Title: Four novel equine papillomaviruses

Contents Category: Animal Viruses - Small DNA

Christian E. Lange^{1,2,*}, Elisabeth Vetsch^{1,2}, Mathias Ackermann¹, Claude Favrot², Kurt
Tobler¹

¹ Institute of Virology, Vetsuisse Faculty, Winterthurerstrasse 266a, CH-8057 Zurich,
Switzerland

² Dermatology Department, Clinic for Small Animal internal Medicine, Vetsuisse Faculty,
Winterthurerstrasse 260, CH-8057 Zurich, Switzerland

Present Address of CEL: Microbiology and Immunobiology Department, Harvard Medical
School, NRB 952, 77 Ave Louis Pasteur, Boston, MA 02115, USA

* Corresponding author

Email: clange@vetclinics.uzh.ch

Tel: +1-617-432-2892

Fax: +1-617-432-4787

Words in summary: 249

Words in main text: 3298

The GeneBank accession numbers for the equine papillomaviruses described are JQ031032
(EcPV4), JQ031033 (EcPV5), JQ965698 (EcPV6) and JX035935 (EcPV7).

Summary

Papillomaviruses appear to be species-specific pathogens, and it was suggested, that each animal species might harbour its own set of papillomaviruses. However all approaches addressing the underlying evolutionary phenomena suffer still from very limited data about animal papillomaviruses. In case of the horse for example only three equine papillomaviruses (EcPVs) have been identified. To further address the situation in this host, supposedly papillomavirus associated lesions were tested for EcPV DNA. Four novel EcPV types were detected and their genomes entirely cloned and sequenced. They display the characteristic organization with early (E) and late (L) regions harbouring the seven classical open reading frames divided by non coding regions. They were named EcPVs 4, 5, 6 and 7, according to their dissimilarity to other papillomaviruses. Most L1 nucleotide identities were shared with EcPV2 in case of EcPV4 (62%) and EcPV5 (60%) or with EcPV3 in case of EcPV6 (70%) and EcPV7 (71%). Thus, EcPVs 4 and 5 may establish novel species within the genus Dyoiota, while EcPVs 6 and 7 might fit into the Dyorho genus and belong to the same species as EcPV3. They were found in genital plaques (EcPV4) aural plaques (EcPV5, EcPV6) or penile masses (EcPV7). Interestingly PCR analysis revealed the DNA of EcPV2 and EcPV4 as well as of EcPV3 and EcPV6 together in the same tissue samples, respectively. In conclusion the DNA of four novel EcPV types was identified and cloned. They cluster with the known types and support broad genetic EcPV diversity in at least two of the known clades. Furthermore, PCR assays also provide evidence for EcPV co-infections in horses.

Introduction

The *Papillomaviridae* family, which has been established as an independent one separated from the *Polyomaviridae* twelve years ago, is rapidly growing, as new putative members are frequently discovered. Their members, the Papillomaviruses (PVs) have a characteristic circular double stranded DNA genome of around 8 kilobase pairs, that usually contains at least six relatively conserved open reading frames in an early (E1, E2, E6, E7) and a late (L1, L2) region. The DNA of the PVs is contained in icosahedral capsids mainly composed of 72 L1 pentamers with a total diameter of approximately 50nm (Howley & Lowy, 2007). The *Papillomaviridae* family is considered as evolutionary very old, as there is a very high genetic diversity among PVs and several animal species have been shown to harbor their own set of PV types. The diversity is revealed in phylogenetic analyses and by 2010 almost 200 PV types had been recognized (Bernard et al., 2010). Almost two thirds of these so far known PV types infect humans, but the number of animal PV discoveries is increasing. The suspected overrepresentation of human PVs in many analyses regarding the phylogeny of PVs has led to different models that are somewhat contradictory (Rector et al., 2007; Bravo et al., 2010). The main obstacle in finding solid ground for a broadly acceptable hypothesis about PV evolution is a very poor resolution outside the human PV genera Alpha, Beta and Gamma.

To classify novel PVs, the DNA sequence of the relatively conserved L1 open reading frame (ORF) is compared with the known ones and accordingly assigned to a genus and species as being a new type, subtype or variant. The criteria for this assignment define genus distinction below 60% identities and species distinction below 70% identities. Novel types have to differ by at least 10% and subtypes by no less than 2%, while smaller differences characterize variants (de Villiers et al., 2004; Bernard et al., 2010). Although this classification does not take into account putative evolutionary differences of the different PV genes or clinical

aspects of PV infections it establishes a viable taxonomic system (de Villiers et al., 2004; Gottschling et al., 2011).

The effects and the impact of PV infections on individuals as well as on public health in general have been partly uncovered in the past decade, with human PVs (HPVs) being in the focus of most efforts. HPVs have been demonstrated to contribute to various benign as well as malignant transformations of epithelial cells, while it was also discovered, that apparently subclinical infections are very frequent (zur Hausen 1996; Antonsson et al., 2000; Howley & Lowy, 2007). Far less is indeed known about PV infections and their effects on animals. Thus, the far best studied animal PV infections include those of domestic animals, which is partly reflected in the number of types known to infect those species, namely 14 in case of dogs, 12 in cattle, 5 in horses (BPV1, BPV2, EcPV1, EcPV2, EcPV3), 3 in sheep and 2 in cats. In this context the horse has a rather prominent role. Although PVs are general regarded as species specific, horses can be infected not only by equine, but also by bovine PVs (BPVs) (Scott & Miller, 2003; Bogaert et al., 2008; Nasir & Campo, 2008). While the infections with BPV1 and BPV2 in horses have been a focus of some research for quite a while, actual equine PVs (EcPVs) were not a big issue until a few years ago. Until now three EcPVs were described, sequenced and cloned in their entity. They all seem to be associated with distinct pathologies. EcPV1 has been isolated from “classical equine papillomas”, a typical benign neoplasia of young horses that is usually transient (O’Banion et al., 1986; Scott & Miller, 2003; Ghin et al., 2004; Postey et al., 2007). In genital neoplasias of benign and malignant character the DNA of EcPV2 has been found and also repeatedly been detected in the most recent years. The data suggest that EcPV2 DNA may indeed be involved in the development of these disorders, although EcPV2 can also be detected in some other equine tumors, in semen and on clinically unaffected skin (Scase et al., 2010; Knight et al., 2011a; Knight et al., 2011b; Lange et al., 2011a; Bogert et al., 2012; Kainzbauer et al., 2012; Sykora et al., 2012; Lange et al., in press). The third one, EcPV3 has been detected in equine aural plaques that had also

previously been suggested to be PV associated (Fairley & Haines, 1994; Postey et al., 2007; Lange et al., 2011a). These three EcPV types do represent also three different PV genera, namely Zeta, Dyoiota and Dyorho. There is also strong evidence for the existence of a fourth EcPV that had been sequenced partially (Taniwaki et al., 2013). The lack of closely related PVs makes a profound phylogenetic allocation of these EcPVs difficult. To improve the knowledge regarding PV phylogeny, to challenge the hypothesis that EcPVs may potentially be as diverse as bovine, canine and maybe human PVs and to find out more about potential involvements in the development of skin disorders this study was initiated.

Results

PCR and RCA Analysis

One of the easiest and most sensitive methods to determine a putative PV association with any type of lesion is classical PCR. Three PCR assays for PV DNA and one for host DNA were therefore applied in all cases included. Host DNA was amplified in all the samples taken (Table 1). In cases #1 (vulval plaques) and #2 (penile and inguinal plaques), DNA was amplified with the FAP59/FAP64 and with the EcPV2probe f/r assays, sequencing revealed EcPV2 DNA as well as the DNA of a thus far unknown PV type. FAP59/FAP64 PCR revealed a different yet also previously not described PV DNA in case #3 (aural plaques). In the samples obtained from case #4 (aural plaques), CP4/CP5 PCR amplified EcPV3 DNA from one ear and a different undetermined PV DNA from the other, while it amplified EcPV3 DNA in both with the WeF03/WeR09 PCR. CP4/CP5 PCR also amplified a novel PV DNA in the samples from case #5 (penile masses). No indication for the presence of HPV DNA was found in any of the cases. Samples containing unknown PV DNA were additionally tested using RCA. The RCA analysis indicated circular DNA of about eight thousand nucleotides in

all these tested samples, whereas the restriction analysis of the RCA products produced patterns unlike any known EcPVs or BPVs. To determine the sequences in detail they were cloned and sequenced. The nucleotide sequences were deposited in GenBank under accession nos. JQ031032 (EcPV4), JQ031033 (EcPV5), JQ965698 (EcPV6) and JX035935 (EcPV7).

Genomes and Open Reading Frames

The genomes of PVs contain usually around 8 kilobasepairs (kbp) and have a genomic organization that typically displays seven or eight open reading frames (ORFs) as well as one or two non-coding-regions (NCRs) (Howley & Lowy, 2007). The genome sizes of the four cloned novel EcPVs were found to be very similar. The genome of EcPV4 consists of 7554 basepairs, EcPV5 of 7519 basepairs, EcPV6 of 7551 basepairs and EcPV7 of 7619 basepairs. The GC contents are 55% for the EcPV4 genome, 50% for the EcPV5 genome, 51% for the EcPV6 genome and 52% for the EcPV7 genome. Seven characteristic ORFs were identified in all four novel EcPV genomes (Fig. 1). The ORFs encoding for the late PV genes appear to be quite conserved regarding their size. The greatest variation among the EcPVs was observed among the early genes E6 and E7. The E4 ORF, which usually lacks an own start codon varies between the viruses, and therefore can not be predicted with certainty.

Characteristic motifs

Several characteristic features on the nucleotide and amino acid level have been proposed and predicted to play a role in the PV life cycle (Androphy et al., 1987; Wilson et al., 2002; Munger et al., 2004; Howley & Lowy, 2007). The genomes of the four novel EcPV candidates were scanned to identify such putative sites and the analysis revealed typical features in all of them (Table 2). Dyad symmetry repeats, a variable number of putative E2

binding sites and at least one poly adenylation signal are present on every EcPV sequence. Sp1 and Nf1 binding sites could be predicted in all EcPVs, AP1 in all but EcPV4 and at least one classical TATA box in all but EcPV2 (Table 2(a)). Examination of the putative protein sequences revealed ATP-dependent helicase motifs in the E1 protein sequence of all EcPVs, but they were modified in the four novel ones. Putative metal-binding motifs were present in all E6 (2) and E7 (1) amino acid sequences. Leucine zipper domains were only found in the E1 of EcPV3 and EcPV7 as well as in the E2 of EcPV2 and EcPV4 (Table 2(b)).

Sequence analysis and comparison

To allocate the novel EcPVs, a phylogenetic analysis based on the aligned E1-E2-L2-L1 sequences was performed. For that purpose the sequences of the eight different EcPV types or isolates as well as forty-two other PVs, representing all presently classified genera were aligned (Fig. 2). The EcPVs were found to cluster together with the Delta PVs with limited support of the branching by the posterior probabilities. The EcPV genera Zeta, Dyoiota and Dyorho were also allocated in relative vicinity of the Chi, Dyotheta and Iota PV genera in this analysis.

For a putative taxonomic classification of the four novel PV isolates according to the present guidelines of PV classification (de Villiers et al., 2004; Bernard et al., 2010), pairwise alignments of the L1 nucleotide sequences were performed. EcPV4 was found to share most identities with EcPV5 (65%) and the two EcPV2 variants (62%). The L1 was also found to be match a partial PV sequence described before (Taniwaki et al., 2013). EcPV5 at the same time apparently shares more identities with EcPV2a (62%) than with EcPV2b (60%). EcPV6 and EcPV7 share most identities with each other (71%) and with EcPV3 (70%). All six obvious ORFs were analyzed in the same manner on the predicted amino acid level (Supplement 1). Over all the highest identities were among the sequences of EcPVs 3, 6 and 7, whereas EcPVs

4 and 5 shared although closest to each other significantly less identities. The only exception was the E6 sequence. In case of E6 the identities between EcPVs 3, 6 and 7 were comparable to those between EcPVs 4 and 5. Consequently EcPV4 and EcPV5 may be regarded as representing two novel species within the genus Dyoiota, whereas EcPV6 and EcPV7 seem to belong to the same species as EcPV3 within the genus Dyorho.

Discussion

The number of known family members within the *Papillomaviridae* is constantly growing. However, while still more than half of the PVs described have been associated with human hosts more and more new discovered animal PVs are starting to alter our view on PV phylogeny and evolution. It appears that a lot if not all host species may harbor their own set of PV types that belong to a few distinct genera each. For the HPVs five such genera are currently recognized containing all of the more than 150 defined types and isolates (de Villiers et al., 2004; Bernard et al., 2010; Kovanda et al., 2011). Analogous to that the bovine and canine PV for example have been assigned to three genera each (Bernard et al., 2010). The same is true for the EcPVs. The here described four novel PV sequences support this hypothesis, as they cluster well together with the few known types within the three defined genera. Upon phylogenetic analysis the three EcPV genera cluster together with each other and with the Delta PVs. This is an interesting observation, as BPV1 and BPV2 that are although classified as bovine PVs also infect horses and cause disease belong to this genus. However, the statistic support for this clade formation is rather weak, and we probably still know only very few (too few) of the circulating animal PVs. Thus drawing conclusion from that would be probably be premature. Seeking for and sequencing further PV candidates may shed more light on these evolutionary and pathobiologically interesting issues.

As the taxonomic classification of PVs is formally based on the L1 nucleotide sequence, the four novel EcPVs could easily be assigned to the relatively novel genera Dyoiota and Dyorho. Besides the genus they may be further classified into species within them. According to their L1 nucleotide sequences EcPV6 and EcPV7 may belong to the same species as EcPV3 within Dyorho. EcPV4 and EcPV5 however, although within Dyoiota like EcPV2 appear to represent two novel species within that genus. This is also supported by the pairwise comparison of the other individual ORFs, that except for E6, showed markedly more identities among the three Dyorho PVs than among the Dyoiota PVs (Supplement 1). The results of the phylogenetic analysis also point in the same direction. The definitive classification however is a matter of the ICTV and ICTV papillomavirus study group respectively. Interestingly there seems to be no correlation between genus and genome size among the EcPVs, whereas that seems to be the case with the HPVs (Alpha>Beta>Gamma), the BPVs (Delta>Epsilon>Xi) and the CPVs (Lambda>Tau>Chi) (Lange et al., 2009).

Previous studies have shown that HPV DNA is not just very abundant on the skin of healthy individuals, but also, that the same host may be infected with or carry more than one PV type at the same time. The results of this study clearly indicate that this is the case in horses as well. In cases #1 and #2 EcPV2 DNA was detected in the same host and the same samples as the novel EcPV4 with different PCR assays. Similarly, in case four EcPV3 and EcPV6 DNA was detected in the same ear with two different the same PCR assays. Only a single double infection different PV types, namely with BPV1 and EcPV2 in a horse and supposedly double infections with different variants of EcPV2 had been described before (Kainzbauer et al., 2012; Lange et al., in press).

One of the most interesting and thus often very controversially discussed questions about PVs is their potential role in the pathogenesis of various neoplastic disorders. The four EcPVs described here have all been extracted from cytobrush samples of lesions that had been

suspected to be PV associated, namely penile neoplasias, aural plaques and depigmented genital and inguinal plaques. Since no biopsy samples could be taken from any of these cases, it is not possible to determine whether any of the four new EcPVs actually played any role in the observed pathologies.

In conclusion four novel EcPVs were discovered and genetically and phylogenetically analysed and characterized. The results support the clustering of EcPVs into the three established genera, confirm infections with more than one PV in horses and warrant further research regarding a possible pathogenicity and involved mechanisms.

Materials and Methods

Samples

For this study, samples of 5 horses were obtained from patients of the veterinary teaching hospital of the Vetsuisse Faculty of the University of Zurich. Inclusion criteria were a suspected diagnosis of PV infection and PCR or RCA results indicating PV DNA of an unknown type. These included one Westfalian mare with white vulval plaques, a Friesian stallion with white inguinal plaques, an Arabian gelding with aural plaques, a German warmblood gelding with aural plaques and an Irish warmblood gelding with penile masses (Table 1). Testing material was acquired with cytobrushes that were moistened with sterile 0.9% NaCl solution, gently rubbed for thirty seconds on putatively PV associated skin alterations and stored in an 1.5ml Eppendorf tube at -20°C until DNA extraction.

Identification, Amplification and cloning of genomes

DNA of the cytobrush samples was isolated using a DNeasy extraction kit (Qiagen). The protocol was slightly modified to meet the specific needs, namely the volume of all buffers used prior to the loading of the column was doubled to ensure the brush to be covered with liquid. The elution buffer was used sparsely (100µl in total) to increase the DNA concentration in the eluate. For initial identification of host and PV DNA in the samples ecGAPDH f (GAG CTG AAT GGG AAG CTC AC) / ecGAPDH r (CTG AGG GCC TTT CTC CTT CT), CP4/CP5, FAP59/FAP64, EcPV2probe f/r and WeF03 (TCT GGG TAG CCG CAA ATA TC) / WeR09 (GCA CCT GAG GTC CTA TCT GG) PCR assays were applied (Forslund et al., 1999; Iftner et al., 2003; Lange et al., in press). The protocols were applied as described, and in case of the ecGAPDH f/r and WeF03/WeR09 primers the same condition as for the EcPV2probe f/r were applied (Lange et al., 2011b; Lange et al., in press). PCR products were determined (Microsynth) by cycle sequencing using an ABI 377 sequencer (Applied Biosystems) and compared to the NCBI database (BLAST X). For the purpose of cloning one microliter of the DNA extract was used for rolling circle amplification (RCA) (Rector et al., 2004), using a TempliPhi Amplification kit (General Electrics Biosciences). Slight modifications were applied to the protocol supplied by the manufacturer (Lange et al., 2009). Amplified PV DNA from cases #1 and #2 (EcPV4) was cloned into the *Bam*HI, *Cla*I, *Eco*RI and *Hin*DIII site of pBluescript II KS+ (Stratagene), into the *Hin*DIII and *Xba*I sites in case of case #3 (EcPV5), into the *Bam*HI site in case of case #4 (EcPV6) and the PV DNA amplified from case #5 was cloned into the *Bam*HI and *Eco*RI sites (EcPV7) using standard procedures. The cloned DNA sequences of the PV candidates were determined (Microsynth) and primary sequences were assembled using Contigexpress software (Vector NTI Informax, Invitrogen). The names for viruses and genera were approved by the papillomavirus study group of the ICTV.

Sequencing analysis

The coding sequences for the E6, E7, E1, E2, L2 and L1 proteins of the four novel EcPVs were compared with previously published PV sequences including: Moose (*Alces alces*): AaPV1 (M15953), bovine (*Bos primigenius*): BPV1 (X02346), BPV3 (NC_004197), BPV5 (AF457465), BPV7 (NC_007612), marine turtles: *Caretta caretta*: CcPV1 (EU493092) and *Chelonia mydas*: CmPV1 (EU493091), caprine (*Capra hircus*): ChPV1 (NC_008032), canine (*Canis lupus familiaris*): CPV1 (L22695), CPV2 (AY722648), CPV3 (NC_008297), CPV4 (NC_010226), equine (*Equus caballus*): EcPV1 (AF498323), EcPV2a (NC_012123), EcPV2b (HM461973), EcPV3 (GU384895), porcupine (*Erethizon dorsatum*): EdPV1 (NC_006951), hedgehog (*Erinaceus europaeus*): EePV1 (EF396272), birds (*Fringilla coelebs*): FcPV1 (NC_004068), *Francolinus leucoscepus*: FIPV1 (EU188799) and *Psittacus erithacus*: PePV1 (NC_003973), feline (*Felis catus*): FdPV1 (NC_004765), FdPV2 (EU796884), human (*Homo sapiens*): HPV1 (NC_001356), HPV4 (NC_001457), HPV5 (NC_001531), HPV9 (NC_001596), HPV16 (FJ610146), HPV18 (NC_001357), HPV41 (NC_001354), HPV50 (NC_001691), HPV63 (NC_001458), hamster: *Mesocricetus auratus*: MaPV1 (E15111), non-human primate (*Macacca mulata*): MmPV1 (NC_001678), muridae (*Micromys minutus*): MmiPV1 (DQ269468) and (*Mastomys natalensis*): MnPV1 (NC_001605), snake (*Morelia spilota*): MsPV1 (HQ262535), rabbit (*Oryctolagus cuniculus*): OcPV1 (NC_002232), Porpoise (*Phocoena spinipinnis*): PsPV1 (NC_003348), fruit bat (*Rousettus aegyptiacus*): RaPV1 (NC_008298), cottontail rabbit (*Sylvilagus floridanus*): SfPV1 (NC_001541), porcine (*Sus scrofa*): SsPV1 (NC_011280), manatee (*Trichechus manatus*): TmPV1 (NC_006563), dolphin (*Tursiops truncatus*): TtPV1 (EU240894), TtPV2 (NC008184) and polar bear (*Ursus maritimus*): UmPV1 (NC_010739). The predicted amino acid sequences were translated and aligned by using MAFFT (Katoh & Toh, 2008) before being back-translated to DNA sequences. The six sets of aligned nucleotide sequences representing the six sets of protein sequences were shortened by using GBLOCK (version 0.91b;

288 <http://molevol.cmima.csic.es/castresana/>) (Castresana 2000) and combined to a single
289 multiple sequence alignment (MSA) by concatenating the sequences from each virus. The
290 optimal model of DNA evolution was evaluated for best fit of the data set using
291 MODELTEST (version 1.4.4; <http://darwin.uvigo.es/>; default settings) (Posada & Crandall,
292 1998). Bayesian phylogeny was inferred using MRBAYES (version 3.2;
293 <http://mrbayes.csit.fsu.edu/>; Markov Chain Monte Carlo with GTR substitution matrix,
294 variable gamma rates, invariant sites, two runs four chains of 10,000,000 generations), and
295 displayed with FIGTREE (1.3.0 <http://tree.bio.ed.ac.uk/>) (Drumond & Rambaut, 2007).
296 Pairwise sequence alignments were performed using NEEDLE (EMBOSS, nucleotide matrix:
297 DNAFULL, amino acid matrix: BLOSSUM62).

298 **Acknowledgements**

299 The authors thank Colin C. Schwarzwald and Meret Wehrli Eser of the Equine Department as
300 well as Nina M. Fischer and Sylvia Wilhelm from the Dermatology Department, Vetsuisse
301 Faculty Zurich for their support with the patients and with clinical aspects and also Melanie
302 A. C. Thumm for support with the sampling.

303

References

- Androphy, E., Lowy, D. & Schiller, J. (1987).** Bovine papillomavirus E2 trans-activating gene product binds to specific sites in papillomavirus DNA. *Nature* **325**, 70-73.
- Antonsson, A., Forslund, O., Ekberg, H., Sterner, G. & Hansson, B. (2000).** The ubiquity and impressive genomic diversity of human skin papillomaviruses suggest a commensalic nature of these viruses. *J Virol* **74**, 11636-11641.
- Bernard, H.-U., Burk, R. D., Chen, Z., van Doorslaer, K., zur Hausen, H. & de Villiers, E. M. (2010).** Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* **401**, 70-79.
- Bogaert, L., Martens, A., Van Poucke, M., Ducatelle, R., De Cock, H., Dewulf, J., De Baere, C., Peelman, L. & Gasthuys, F. (2008).** High prevalence of bovine papillomaviral DNA in the normal skin of equine sarcoid-affected and healthy horses. *Vet Microbiol* **129**, 58-68.
- Bogert, L., Willemsen, A., Vanderstraeten, E., Bracho, M. A., De Baere, C., Bravo, I. G. & Martens, A. (2012).** EcPV2 DNA in equine squamous cell carcinomas and normal genital mucosa. *Vet Microbiol* **158**, 33-41.
- Bravo, I. G., de Sanjosé, S. & Gottschling, M. (2010).** The clinical importance of understanding the evolution of papillomaviruses. *Trends Microbiol* **18**, 432-438.
- Castresana, J. (2000).** Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* **17**, 540-552.
- de Villiers, E. M., Fauquet, C., Broker, T. R., Bernard, H.-U. & zur Hausen, H. (2004).** Classification of papillomaviruses. *Virology* **324**, 17-27.
- Drummond, A. & Rambaut, A. (2007).** BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**, 214.
- Fairley, R. & Haines, D. (1994).** The electron microscopic and immunohistochemical demonstration of a papillomavirus in equine aural plaques. *Vet Pathol* **29**, 79-81.

329 **Forslund, O., Antonsson, A., Nordin, P., Stenquist, B. & Hansson, B. (1999).** A broad
330 range of human papillomavirus types detected with a general PCR method suitable of
331 cutaneous tumours and normal skin. *J Gen Virol* **80**, 2437-2443.

332 **Ghim, S., Rector, A., Delius, H., Sundberg, J., Jenson, A. & Van Ranst, M. (2004).**
333 Equine papillomavirus type 1: complete nucleotide sequence and characterization of
334 recombinant virus-like particles composed of the EcPV-1 L1 major capsid protein. *Biochem*
335 *Biophys Res Commun* **324**, 1108-1115.

336 **Gottschling, M., Göker, M., Stamatakis, A., Bininda-Emonds, O. R. P., Nindl, I. &**
337 **Bravo, I. G. (2011).** Quantifying the phylodynamic forces driving papillomavirus evolution.
338 *Mol Biol Evo* **28**, 2101-2113.

339 **Iftner, A., Klug, S., Garbe, C., Blum, A., Stancu, A., Wilczynski, S. & Iftner, T. (2003).**
340 The prevalence of human papillomavirus genotypes in nonmelanoma skin cancers of
341 nonimmunosuppressed individuals identifies high-risk genital types as possible risk factors.
342 *Cancer Res* **63**, 7515-7519.

343 **Howley, P. M. & Lowy, D. R. (2007).** Papillomaviruses. In *Fields Virology Volume 2*. 5th
344 Edition, pp. 2299-2354. Edited by Knipe, D. M. & Howley, P. M. Philadelphia, PA, USA:
345 Lippincott Williams and Wilkins.

346 **Kainzbauer C, Rushton J, Tober R, Scase T, Nell B, Sykora S, Brandt S. (2012).** Bovine
347 papillomavirus type 1 and Equus caballus papillomavirus 2 in equine squamous cell
348 carcinoma of the head and neck in a Connemara mare. *Equine Vet J* **44**, 112-115.

349 **Katoh, K. & Toh, H. (2008).** Recent developments in the MAFFT multiple sequence
350 alignment program. *Brief Bioinform* **9**, 286-298.

351 **Knight, C. G., Munday, J. S., Peters, J. & Dunowska, M. (2011).** Equine Penile Squamous
352 Cell Carcinomas Are Associated With the Presence of Equine Papillomavirus Type 2 DNA
353 Sequences. *Vet Pathol* **48**, 1190-1194.

354 **Knight, C. G., Munday, J. S., Rosa, B. V. & Kiupel, M. (2011).** Persistent, widespread
355 papilloma formation on the penis of a horse: a novel presentation of equine papillomavirus
356 type 2 infection. *Vet Dermatol* **22**, 570-574.

357 **Kovanda, A., Kocjan, B. J., Luzar, B., Bravo, I. G. & Poljak, M. (2011).** Characterisation
358 of novel cutaneous human papillomavirus genotypes HPV-150 and HPV-151. *PLoS One* **6**,
359 e22529.

360 **Lange, C. E., Tobler, K., Ackermann, M., Panakova, L., Thoday, K. L. & Favrot, C.**
361 **(2009).** Three novel canine papillomaviruses support taxonomic clade formation. *J Gen*
362 *Virology* **90**, 2615-2621.

363 **Lange, C. E., Tobler, K., Ackermann, M. & Favrot, C. (2011a).** Identification of two novel
364 equine papillomavirus sequences suggests three genera in one cluster. *Vet Microbiol* **149**, 85-
365 90.

366 **Lange, C. E., Zollinger, S., Tobler, K., Ackermann, M. & Favrot, C. (2011b).** The
367 clinically healthy skin of dogs is a potential reservoir for canine papillomaviruses. *J Clin*
368 *Microbiol* **49**, 707-709.

369 **Lange, C. E., Tobler, K., Lehner, A., Grest, P., Welle, M., Schwarzwald, C. C. & Favrot,**
370 **C. (in press).** EcPV2 DNA in Equine Papillomas, In Situ and Invasive Squamous Cell
371 Carcinomas Supports Papillomavirus Etiology. *Vet Pathol*.

372 **Münger, K., Baldwin, A., Edwards, E., Hayakawa, H., Nguyen, C., Owens, M., Grace,**
373 **M. & Huh, K. (2004).** Mechanisms of human papillomavirus-induced oncogenesis. *J Virol*
374 **78**, 11451-11460.

375 **Nasir, L. & Campo, M. (2008).** Bovine papillomaviruses: their role in the aetiology of
376 cutaneous tumours of bovids and equids. *Vet Dermatol* **19**, 243-254.

377 **O'Banion, M., Reichmann, M. & Sundberg, J. (1986).** Cloning and characterization of an
378 equine cutaneous papillomavirus. *Virology* **152**, 100-109.

379 **Posada, D. & Crandall, K. (1998).** MODELTEST: testing the model of DNA substitution.
380 *Bioinformatics* **14**, 817-818.

381 **Rector, A., Tachezy, R. & Van Ranst, M. (2004).** A sequence-independent strategy for
382 detection and cloning of circular DNA virus genomes by using multiply primed rolling-circle
383 amplification. *J Virol* **78**, 4993-4998.

384 **Rector, A., Lemey, P., Tchetzy, R., Mostmans, S., Ghim, S.-J., Van Doorslaer, K.,**
385 **Roelke, M., Bush, M., Montali, R. J., Joslin, J., Burk, R. D., Jenson, A. B., Sundberg, J.**
386 **P., Shapiro, B. & Van Ranst, M. (2007).** Ancient papillomavirus-host co-speciation in
387 Felidae. *Genome Biol* **8**, R57.

388 **Postey, R. C., Appleyard, G. D. & Kidney B. A. (2007).** Evaluation of equine papillomas,
389 aural plaques, and sarcoids for the presence of Equine papillomavirus DNA and
390 Papillomavirus antigen. *Can J Vet Res* **71**, 28-32.

391 **Scase, T., Brandt, S., Kainzbauer, C., Sykora, S., Bijmolt, S., Hugher, K., Sharpe, S. &**
392 **Foote, A. (2010).** Equus caballus papillomavirus-2 (EcPV-2): An infectious cause for equine
393 genital cancer? *Equine Vet J* **42**, 738-745.

394 **Scott, D. W. & Miller Jr, W. H. (2003).** Equine Dermatology, pp700-712. St. Louis, MO,
395 USA: Saunders.

396 **Sykora, S., Samek, L., Schönthaler, K., Palm, F., Borzacchiello, G., Aurich, C. &**
397 **Brandt, S. (2012).** EcPV-2 is transcriptionally active in equine SCC but only rarely detectable
398 in swabs and semen from healthy horses. *Vet Microbiol* **158**, 194-198.

399 **Taniwaki, S. A., Magro, A. J., Gorino, A. C., Oliveira-Filho, J. P., Fontes, M. R. M.,**
400 **Borges, A. S. & Araujo Jr, J. P. (2013).** Phylogenetic and structural studies of a novel
401 equine papillomavirus identified from aural plaques. *Vet Microbiol* **162**, 85-93.

402 **Wilson, V., West, M., Woytek, K. & Rangasamy, D. (2002) Papillomavirus E1 proteins:**
403 **form, function, and features.** *Virus Genes* **24**, 275-290.

404 **Zur Hausen, H. (1996).** Papillomavirus infections – a major cause of human cancers.
405 *Biochim Biophys Acta* **1288**, F55-F78.
406

406 **Figure legends**

407 Fig. 1

408 Schematic genomes of equine papillomaviruses

409 Schematic presentation of the seven equine papillomavirus genomes and their open reading
410 frames (ORFs). Genomes are divided into sections: early genes (Early), late genes (Late) and
411 non-coding regions (NCRs). Numbers indicate nucleotide positions. Nucleotide position
412 number one is defined here as the first following the stop codon of the L1 ORFs.

413

414 Fig. 2

415 Phylogenetic tree of 50 papillomaviruses

416 Midpoint rooted Bayesian phylogenetic tree of 50 papillomaviruses (PVs) representing most
417 genera. Numbers at internal nodes represent the posterior probability support values.

418 Unclassified PVs are marked with asterisks. Equine PVs are marked red.

419

419 **Table 1:** Included cases

420

Case No.	Gender	Breed	Age	Symptoms	Samples of	results of PCR sequencing
1	Mare	Westfalian	6	White vulval plaques	Plaque	EcPV2, EcPV4
2	Gelding	Frisian	10	White penile and inguinal plaques	Penis	EcPV2
					Preputium	EcPV2
					Inter inguinal	EcPV2, EcPV4
3	Gelding	Arabian	5	Aural plaques	Plaque	EcPV5
4	Gelding	German warm blood	8	Aural plaques left ear	Left ear 's rim	EcPV3
					Central left pinna	EcPV3
					Central right pinna	EcPV3, EcPV6
5	Gelding	Irish warm blood	22	Penile masses	Penile mass	EcPV7

421

422

Table 2

Characteristic features on nucleotide and amino acid level

(a)

Predicted nt-feature	EcPV1	EcPV2	EcPV3	EcPV4	EcPV5	EcPV6	EcPV7
E2 binding site (ACC-N ₅₋₇ -GGT)	116; 3028; 4168; 5298; 5388; 6527; 7144; 7378; 7509	318; 4828; 5193; 7591; 7760	358; 650; 1148; 4142; 4226; 5001; 6988; 7004; 7124; 7218; 7310; 7337; 7468	811; 1458; 7133	6971; 7106; 7150	1643; 4062; 6986; 7033; 7245; 7482	4276; 7001; 7024; 7047; 7136; 7243; 7264; 7513
Dyad symmetry repeats (TTGTTGTAAACAACAA)	7550	145	7510	7475	7453	7524	7555; 3693
Poly adanylation sites (AATAAA)	2255; 3749; 3756; 6882	7216	2345; 6748; 6808	3937; 6062; 6944	6928	2381; 2444; 5785; 6397; 6845	2600; 6273; 6869
Sp1 binding sites (GGCGGG)	3815; 3997; 4558; 5982; 7357	4707; 4929; 5180; 6723; 7788	1915; 4272; 4311; 4798; 6967	1077; 2062; 4949; 5352; 6168; 7545	1170; 3310; 3535; 7282	1207; 3878; 4848;	41; 1196; 3675; 4585;
Nf1 binding sites (CGGAA)	483; 3329	2906; 3304; 3421; 3575; 3895; 5435	361; 687; 2946; 5591	937; 1170; 1764; 1773; 3035; 3258; 5531; 5612; 5883; 6366; 6890; 7120; 7310; 7361; 7412; 7442	2639; 2698; 2792; 2987; 6346; 7254; 7389;	466; 1830; 2139; 2280; 2737; 5006; 5977; 6046; 6175; 6394; 7059; 7275;	613; 1304; 1791; 1828; 2143; 2278; 2707; 2766; 3940; 4416; 5997; 6417; 7130; 7307; 7351
API binding site (TGANTCA)	3390; 4902	97	7459	-	7407	700; 7068; 7197; 7473	683; 7187
Tata signals (TATAAA or TATA(A/T)A(A/T))	6049; 6051; 7600	-	1563; 6547; 6827	5576	6376	276; 5488; 6357; 6946	1597; 6961

427 (b)

Predicted aa-feature	EcPV1	EcPV2	EcPV3	EcPV4	EcPV5	EcPV6	EcPV7
ATP-dependent helicase motifs in E1 (GPPNTGKS)	443aa/2016nt	455aa/2305nt	443aa/2028nt	453aa/2007nt T for S	458aa/2031 D for N	438aa/2061nt D for N	444aa/2061nt D for N
metal-binding motifs in E6 (CX ₂ CX ₂₈₋₃₀ CX ₂ C)	12aa/58nt; 85aa/277nt	8aa/260; 83aa/485nt	21aa/64nt; 94aa/283nt	10aa/58nt; 84aa/280nt	10aa/70nt; 84aa/292nt	19aa/114nt; 92aa/331nt	23aa/97nt; 96aa/316nt
metal-binding motifs in E7 (CX ₂ CX ₂₈₋₃₀ CX ₂ C)	53aa/508nt	70aa/813nt	47aa/572nt	40aa/518nt	39aa/530nt	45aa/620nt	45aa/603nt
Leucine zipper domain in E1 (LX ₅₋₇ LX ₅₋₇ LX ₅₋₇ L)	-	-	264aa/1491nt	-	-	268aa/1552nt	273aa/1549nt
Leucine zipper domain in E2 (LX ₅₋₇ LX ₅₋₇ LX ₅₋₇ L)	-	4aa/2777nt	-	4aa/2479nt	-	-	-

428

429



